Prof. Dr. Rupert Lasser
Paul Bergold

Technische Universität München
Zentrum Mathematik

Stochastik für Lehramt Gymnasium – Blatt 7

Wintersemester 2016/17

Die Abgabe zu Blatt 7 erfolgt in der Woche vom **12.12. bis 16.12.2016.**
Fragen und Hinweise bitte an bergold@ma.tum.de.

Übungen (Vergleich von Merkmalen, RStudio)

Aufgabe 1

Gegeben sei folgende Messreihe (x_i, y_i) in Abhängigkeit von $u \in \mathbb{R}$:

i	1	2	3	4	5
x_i	1	3	7	9	5
y_i	2	4	u	10	6

1. Bestimmen Sie u so, dass $r_{XY} = 1$ bzw. $r_{XY} = 0$. Begründen Sie Ihre Antwort.
2. Sei nun $u = 6$. Bestimmen Sie die Gleichung der Trendgeraden $R_{Y(X)}$.
3. Zeichnen Sie den Punktschwarm zu den Messdaten und zeichnen Sie die Trendgerade ein.

Aufgabe 2

Gegeben seien die Werte $x_1, x_2, ..., x_m$ und $y_1, y_2, ..., y_m$ zweier Merkmale X und Y. Gesucht sind Koeffizienten a und b, welche die Gleichungen

$$\begin{aligned} ax_1 + b &= y_1, \\ ax_2 + b &= y_2, \\ &\vdots \\ ax_m + b &= y_m, \end{aligned} \tag{1}$$

erfüllen. Da dieses System im Allgemeinen keine Lösung besitzt, wählt man a und b so, dass die Residuen $r_i(a, b) := y_i - b - ax_i$ möglichst klein sind. Der Ansatz der Methode der kleinsten Quadrate lautet:

$$\text{Minimiere } \sum_{i=1}^{m} (y_i - b - ax_i)^2. \tag{2}$$

1. Schreiben Sie das Gleichungssystem (1) in der Form $Av = y$ mit einer Matrix $A \in \mathbb{R}^{m \times 2}$, $v := (a, b)^T \in \mathbb{R}^2$ und $y \in \mathbb{R}^m$.

2. Das Minimierungsproblem (2) kann wie folgt geschrieben werden:

$$\min_{v \in \mathbb{R}^2} f(v).$$

   Dabei ist die Zielfunktion $f\colon \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(v) = \|y - Av\|_2^2$. Bestimmen Sie $\nabla f(v)$.

3. Ist v^* ein Minimum der Funktion f, so muss $\nabla f(v^*) = 0$ erfüllt sein. Leiten Sie die Normalgleichung $A^T A v^* = A^T y$ her.

Hausaufgaben

Hausaufgabe 19 (Diese Hausaufgabe erstellen Sie bitte mit RStudio)
Gegeben sei die folgende Übersicht der Ergebnisse einer Schulaufgabe. Es liegt das Bewertungsschema von 1 (sehr gut) bis 6 (ungenügend) zugrunde. Zudem sind noch Mitarbeitsnoten gegeben:

Schüler	1	2	3	4	5	6	7	8	9	10
Note Schulaufgabe (Y)	2	3	1	2	5	3	2	2	4	5
Note Mitarbeit (X)	2	2	1	1	3	2	2	1	3	3

1. Erstellen Sie die Vektoren `x` (Note Mitarbeit) und `y` (Note Schulaufgabe).

2. Zeichnen Sie die empirische Verteilungsfunktion und einen Boxplot für Y.

3. Berechnen Sie den Korrelationskoeffizienten.

4. Verwenden Sie den Befehl `lm(y~x)`, um die Gleichung der Trendgeraden $R_{Y(X)}$ zu bestimmen.

5. Tragen Sie die obigen Werte als Punktwolke in ein Koordinatensystem ein und fügen Sie anschließend die Regressionsgerade hinzu.

6. Im Raum steht die Behauptung, dass die Mitarbeitsnoten und die Ergebnisse der Schulaufgabe stark voneinander abweichen. Nehmen Sie aufgrund Ihrer Ergebnisse kritisch Stellung zu dieser Aussage.

Hausaufgabe 20 (Diese Hausaufgabe erstellen Sie bitte mit RStudio)
Betrachten Sie die Merkmale X und Y aus Hausaufgabe 19. Ziel ist es, mit Hilfe der Normalgleichung aus Aufgabe 2 die Gleichung der Trendgeraden $R_{Y(X)}$ zu berechnen.

1. Erzeugen Sie die Matrix A aus Aufgabe 2.

2. Berechnen Sie $B := A^T A$ und $w := A^T y$

3. Zeigen Sie, dass $\text{rang}(B) = 2$ und somit die Normalgleichung eindeutig lösbar ist.

4. Lösen Sie die Normalgleichung und geben sie v^* an. Wie lautet die Gleichung der Trendgeraden?

Hausaufgabe 21

Gegeben sei ein Merkmal $X\colon M \to [4], \#M = 240$, durch folgende Tabelle:

a_i	1	2	3	4
$r(X = a_i)$	$c/2$	$c/4$	$c/8$	$c/16$

1. Bestimmen Sie den Wert von c.

2. Sei $Y\colon M \to \mathbb{R}$ ein weiteres Merkmal, gegeben durch:

b_i	5	6	7	8
$r(Y = b_i)$	$8/15$	$4/15$	$2/15$	$1/15$

   Wie müssen a und b gewählt werden, dass $Y = aX + b$ gilt?

3. Begründen Sie (ohne explizites Ausrechnen), dass $\bar{y} = \bar{x} + 4$ und $\sigma_X^2 = \sigma_Y^2$.

4. Wie viele Einträge besitzt δ_X?

5. Geben Sie die Gleichungen der Trendgeraden $R_{Y(X)}$ und $R_{X(Y)}$ an.